

WEST Search History

DATE: Monday, November 04, 2002

<u>Set Name</u> side by side	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u> result set
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR</i>			
L25	l3 and l20	9	L25
<i>DB=USPT; PLUR=YES; OP=OR</i>			
L24	5987117.pn.	1	L24
L23	6014647.pn.	1	L23
L22	6044146.pn.	1	L22
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR</i>			
L21	l13 and L20	3	L21
L20	L19 and meta same information	41	L20
L19	L18 and multiple same party	247	L19
L18	search same engines	7049	L18
L17	L16 and multiple same party	1	L17
L16	meta adj search adj engines	40	L16
L15	((709/245)!.CCLS.))	668	L15
L14	((709/227)!.CCLS.))	1180	L14
L13	((709/224)!.CCLS.))	1652	L13
L12	((709/218)!.CCLS.))	852	L12
L11	((345/968)!.CCLS.))	112	L11
L10	((345/866)!.CCLS.))	293	L10
L9	((345/\$)!.CCLS.))	48196	L9
L8	((707/10)!.CCLS.))	2468	L8
L7	((707/7)!.CCLS.))	547	L7
L6	((707/6)!.CCLS.))	790	L6
L5	((707/4)!.CCLS.))	1113	L5
L4	((707/2)!.CCLS.))	1194	L4
L3	((707/\$)!.CCLS.))	17052	L3
L2	((707/5)!.CCLS.))	969	L2
L1	((707/3)!.CCLS.))	2197	L1

END OF SEARCH HISTORY

4 citations found. Retrieving documents...

Quarterdeck, Inc. *Webcompass Fact Sheet*.

<http://www.arachnid.qdeck.com/qdeck/products/webcompass>.

CiteSeer [Home/Search](#) [Document Not in Database](#) [Summary](#) [Related Articles](#) [Check](#)

This paper is cited in the following contexts:

A Schema-based Approach to Modeling and Querying WWW Data - Comai, Damiani, Posenato.. (1998)
(Correct)

....search engine relies on a broad hierarchical classification systems of subjects, much similar to those used by the Library of Congress. **Yahoo s success has spawned multiple similar tools, all based on the idea of providing large, monolithic servers holding indexes of site contents (WebCompass [Qua] SavvySearch [Dre] and others) 3 Structural representation of sites.** A considerable amount of research has been made on how to provide database style support for querying the Web, and three main WWW query languages have been proposed so far: Web3QL [KS95] WebSQL [MMM96] and WebLog [LSS96]

....Local Query Manager presentation issues Instance Query Answer Query partial answer Client Area User Schema Search Area (Robot) Interface Query (graphical) Schema Robot Query Manager Remote Figure 1: Outline of our system architecture. **tion. Since many current Web users are well acquainted with graph like representation of the hypertextual structure of Web sites, the main purpose of our client module is to provide user friendly visual tools for query formulation, based on graph like schema and query representation.** Novice users can rely on a cut and paste approach to put

[Article contains additional citation context not shown here]

Quarterdeck, Inc. *Webcompass Fact Sheet*.

<http://www.arachnid.qdeck.com/qdeck/products/webcompass>.

The WG-Log System: Data Model and Semantics - Comai (1998) (Correct)

....search engine relies on a broad hierarchical classification systems of subjects, much similar to those used by the Library of Congress. **Yahoo s success has spawned multiple similar tools, all based on the idea of providing large, monolithic servers holding indexes of site contents (WebCompass [Qua97], SavvySearch [Dre97] and others) Structural representation of sites.** A considerable amount of research has been made on how to provide database style support for querying the Web, and three main WWW query languages have been proposed so far: Web3QL [Kon95] WebSQL [Men96] and WebLog [Lak96]

Quarterdeck Inc. "*Web Compass Fact Sheet*"

<http://www.arachnid.qdeck.com/qdeck/products/webcompass/>

Semantic Approaches to Structuring and Querying Web Sites - Damiani, Tanca (1997) (1 citation)
(Correct)

.... text searching, obtaining the power of a full Related Work 7 Basato Figure 3 A semantics based classification of research approaches to WWW querying 8 Semantic Approaches to Structuring and Querying Web Sites fledged text retrieval system, Meta search services (MetaCrawler [Sel95]

WebCompass [Qua97], SavvySearch [Dre97] and others) have been built that are able to use powerful free text indexes like Altavista [Alt97] as subroutines, querying all available services in parallel and then aggregating the results. Although these search engines present a sophisticated query interface, the results

Quarterdeck Inc. "*Web Compass Fact Sheet*"

<http://www.arachnid.qdeck.com/qdeck/products/webcompass/>

Structuring and Querying the Web through Graph-Oriented Languages - Damiani (1997) (2 citations)
(Correct)

....ACM Classification of Computer Science Topics. **Yahoo s success has spawned multiple similar tools, all based on the idea of providing large, monolithic servers holding indexes of site contents (Point [Lyc97 2] Magellan [McK97] Recently, metasearch services (MetaCrawler [Sel95] WebCompass [Qua97], SavvySearch [Dre97] and others) have been built that are able to use monolithic indexes as subroutines, querying all available services in parallel and then aggregating the results.** Although many of these search engines present a sophisticated query interface, the results they deliver are

Quarterdeck, Inc., "*Webcompass Fact Sheet*"

<http://www.arachnid.qdeck.com/qdeck/products/webcompass/> 14

[Online articles have much greater impact](#) [More about CiteSeer](#) [Add search form to your site](#) [Submit documents](#) [Feedback](#)
[Latest news](#)

CiteSeer - citeseer.org - [Terms of Service](#) - [Privacy Policy](#) - Copyright © 1997-2002 [NEC Research Institute](#)

[Advanced Search](#) [Preferences](#) [Language Tools](#) [Search Tips](#)

web compass

Google Search

[Web](#) [Images](#) [Groups](#) [Directory](#) [News](#)Searched the web for **web compass**.Results **111 - 120** of about **471,000**. Search took **0.35** seconds.**COMPASSES - LOWEST Prices at DealTime - Click HERE!**

Sponsored Link

www.DealTime.com Compare Prices from 1,000s of Stores - Find the Best Deals HERE!

Compass - Search for Sports Items on eBay

Sponsored Link

www.eBay.com All Sports Play Here! - eBay Sports

Lonergan Web Site: Republication of Compass... Lonergan **Web** Site thanks the Special Issue Editors, Frederick E. Crowe, SJ and Robert M. Doran, SJ for their kind permission to republish the **Compass** journal ...www.lonergan.on.ca/reprints/compass.htm - 6k - [Cached](#) - [Similar pages](#)**6cm with Web & Compass**View Cart. 6cm with **Web & Compass**, \$3.00. A great 6cm carabiner with **web** fob and **compass**. Color: Blue. Quantity: ...www.bisondesigns.com/Catalog/ 6cm_with_web___compass_661214.htm - 44k - [Cached](#) - [Similar pages](#)[[More results from www.bisondesigns.com](#)]**appointment scheduling for doctors, physicians and hospitals - ...**... If your clients already use an existing practice management program, we may be able to integrate **COMPASS** with that program to create a **web**-based solution. ...www.compasssscheduling.com/resellers.htm - [Similar pages](#)[[More results from www.compasssscheduling.com](#)]**Compass West - International School of Rock Climbing**Welcome to the **web** site of **Compass** West. We are a small family run organisation dedicated to providing adventure and skills training ...www.compasswest.co.uk/ - 11k - [Cached](#) - [Similar pages](#)**Compass Announces Reports-To-Web**... Inc. by contacting 800-832-7021. For more product information, see the **Compass web** site at www.compasscs.com. About Clovis & Roche: ...

www.compasscs.com/Web%20Collaterals/

PR%20Compass%20Moves%20Clovis%20and%20Roche%20into%2021st%20Century.htm - 14k - [Cached](#) -[Similar pages](#)**The Location Toolbar Your Compass To The Web - Smart Computing**... solutions, hardware answers, PC basics, software solutions, software tutorials, hardware advice. Featured : The Location Toolbar - Your **Compass** To The **Web**. ...www.smartcomputing.com/editcat/SMART/ SPECIFIC+SOFTWARE+PRODUCTS/221/1987/ - [Similar pages](#)**Compass Consulting Group: Web Demos**www.compasscg.com/web_demos/ - 1k - [Cached](#) - [Similar pages](#)**Caribbean Compass Homepage**... Vincent & the Grenadines Website designed by **Compass** Design. If you would like a **web** presence in St Vincent & the Grenadines (or indeed, anywhere!), contact us ...

Description: Free monthly newspaper covering marine and shore based topics in the Southern Caribbean from Martinique...

Category: [Regional](#) > [Caribbean](#) > ... > [Newspapers and Publications](#)

[www.caribbeancompass.com/](#) - 11k - Nov. 2, 2002 - [Cached](#) - [Similar pages](#)

Compass Informatics - Web Design

Web Design & Hosting. **Compass** Informatics are established developers of innovative quality **web** sites. In addition we also provide full site hosting services. ...

[www.compass.ie/webdesign/main.html](#) - 6k - [Cached](#) - [Similar pages](#)

Compass Informatics - Web & MultiMedia Services

Web and MultiMedia Solutions. **Compass** Informatics Ltd. 1997.

[www.compass.ie/webmedia/main.html](#) - 2k - [Cached](#) - [Similar pages](#)

[[More results from www.compass.ie](#)]

◀ Goooooooooooooooooooooooooooooogle ▶

Result Page: [Previous](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#) [16](#) [17](#) [18](#) [19](#) [20](#) [21](#) [Next](#)

[Search within results](#)

[Google Home](#) - [Advertise with Us](#) - [Search Solutions](#) - [Language Tools](#) - [Jobs, Press, & Help](#)

©2002 Google

[> home](#) [> about](#) [> feedback](#) [> logout](#)

US Patent & Trademark Office

Search Results





Search Results for: [web <and> compass]

Found 134 of 102,695 searched. → Rerun within the Portal







Search within Results

multiple <and> party <and> search <and> engines[> Advanced Search](#) [> Search Help/Tips](#)**Sort by:** Title Publication Publication Date Score Binder**Results 1 - 20 of 134** short listing[1](#)[2](#)[3](#)[4](#)[5](#)[6](#)[7](#)

- 1** Shared visiting in EQUATOR city 84%
 Ian MacColl , Dave Millard , Cliff Randell , Anthony Steed , Barry Brown , Steve Benford , Matthew Chalmers , Ruth Conroy , Nick Dalton , Areti Galani , Chris Greenhalgh , Danius Michaelides , Tom Rodden , Ian Taylor , Mark Weal
Proceedings of the 4th international conference on Collaborative virtual environments September 2002
In this paper we describe an infrastructure and prototype system for sharing of visiting experiences across multiple media. The prototype supports synchronous co-visiting by physical and digital visitors, with digital access via either the World Wide Web or 3-dimensional graphics.
- 2** Using singular value decomposition to visualise relations within multi-agent systems 82%
 Michael Schroeder
Proceedings of the third annual conference on Autonomous Agents April 1999
- 3** Quest for Java 82%
 Sonya Freeman Cohen
Communications of the ACM January 1998
Volume 41 Issue 1

- 4** Next-Gen Open Hypermedia, Part One: Towards geo-spatial hypermedia: Concepts and prototype implementation 80%
 Kaj Grønbaek , Peter Posselt Vestergaard , Peter Ørbæk
Proceedings of the thirteenth conference on Hypertext and hypermedia June 2002
This paper combines spatial hypermedia with techniques from Geographical Information Systems and location based services. We describe the Topos 3D Spatial Hypermedia system and how it has been developed to support geo-spatial hypermedia coupling hypermedia information to model representations of real world buildings and landscapes. The prototype experiments are primarily aimed at supporting architects and landscape architects in their work on site. Here it is useful to be able to superimpose and ...
- 5** Enterprise resource planning: making ERP a success 80%
 August-Wilhelm Scheer , Frank Habermann
Communications of the ACM April 2000
Volume 43 Issue 4
- 6** Pen computing: a technology overview and a vision 80%
 André Meyer
ACM SIGCHI Bulletin July 1995
Volume 27 Issue 3
This work gives an overview of a new technology that is attracting growing interest in public as well as in the computer industry itself. The visible difference from other technologies is in the use of a pen or pencil as the primary means of interaction between a user and a machine, picking up the familiar pen and paper interface metaphor. From this follows a set of consequences that will be analyzed and put into context with other emerging technologies and visions. Starting with a short historic ...
- 7** Student Posters: Backseat gaming: expolaration of mobile properties for fun 80%
 Liselott Brunnberg
Conference Extended Abstracts on Human Factors in Computer Systems April 2002
This paper presents a prototype developed as part of the Backseat gaming project. The aim of the project is to explore how to make use of moblie properties for developing compelling and fun game experiences. The prototype is developed for use in a highly mobile situation, that of a car passenger and is realized by the use of mobile devices and the users physical location during speed to

merge the virtual content and surrounding road context into an augmented reality game.

- 8** Semiautomatic generation of glossary links: a practical solution 80%
 Hermann Kaindl , Stefan Kramer , Papa Samba Niang Diallo
Proceedings of the tenth ACM Conference on Hypertext and hypermedia : returning to our diverse roots: returning to our diverse roots February 1999
- 9** Business designs of the university 80%
 Peter Denning
ACM Computing Surveys (CSUR) December 1996
- 10** Bibliography of recent publications on computer communication 80%
 Martha Sreenstrup
ACM SIGCOMM Computer Communication Review October 1995
Volume 25 Issue 5
- 11** Local and global properties in networks of processors (Extended Abstract) 77%
 Dana Angluin
Proceedings of the twelfth annual ACM symposium on Theory of computing April 1980
This paper attempts to get at some of the fundamental properties of distributed computing by means of the following question: "How much does each processor in a network of processors need to know about its own identity, the identities of other processors, and the underlying connection network in order for the network to be able to carry out useful functions?" The approach we take is to require that the processors be designed without any knowledge (or only very broad knowledge) o ...
- 12** Work structures and shifts: an empirical analysis of software specification teamwork 77%
 Salah Bendifallah , Walt Scacchi
Proceedings of the 11th international conference on Software engineering May 1989
- 13** The relationship between business and higher education: a perspective on the 21st century 77%
 John Sculley
Communications of the ACM September 1989
Volume 32 Issue 9

Apple's Chairman of the Board discusses the future of three core technologies—hypermedia, simulation and artificial intelligence—and the role each will play in education. The following speech was presented to an audience of teachers almost two years ago. Its message, however, is as timely and inspirational today as we prepare for a new era.

14 Work group structures and computer support: a field experiment 77%

 J. D. Eveland , Tora K. Bikson

Proceedings of the conference on Computer-supported cooperative work January 1988

What happens when task groups attempt to couple the advantages of online text preparation or data analysis and decision support with computer-based communication capabilities? How, if at all, does networked information technology affect group structures and interaction processes? And do positive answers to these questions depend on having a technology rich environment with computer-sophisticated individuals to start with, or could almost anyone reap significant advantages if provided with b ...

15 Work group structures and computer support: a field experiment 77%

 J. D. Eveland , T. K. Bikson

ACM Transactions on Information Systems (TOIS) October 1988
Volume 6 Issue 4

It is frequently suggested that work groups that have computer technology to support activities such as text editing, data manipulation, and communication develop systematically different structures and working processes from groups that rely on more conventional technologies such as memos, phone calls, and meetings. However, cross-sectional or retrospective research designs do not allow this hypothesis to be tested with much power. This field experiment created two task forces, each compos ...


16 The Silicon Valley Sentinel-Observer: In-depth!: the 77%

 Sentinel-Observer's public affairs television program

Richard G. Epstein

ACM SIGCAS Computers and Society March 2000
Volume 30 Issue 1

17 Short papers: Developing 3D information systems for mobile 77%


 users: some usability issues

Teija Vainio , Outi Kotala

Proceedings of the second Nordic conference on Human-computer interaction October 2002

Location-based services, navigation and way finding in three-dimensional worlds are some challenges of current research in mobile computing and information systems. This paper introduces a prototype of future mobile city information systems. In 3D City Info research project, we integrated in one prototype system a three-dimensional city model, a map and a database, which includes information from the same area. In addition, we have built a fully working mobile laptop version of the 3D City Info ...


18 Papers: novel input, output, and computation: TiltType: 77%

 accelerometer-supported text entry for very small devices
Kurt Partridge , Saurav Chatterjee , Vibha Sazawal , Gaetano Borriello , Roy Want

Proceedings of the 15th annual ACM symposium on User interface software and technology October 2002


TiltType is a novel text entry technique for mobile devices. To enter a character, the user tilts the device and presses one or more buttons. The character chosen depends on the button pressed, the direction of tilt, and the angle of tilt. TiltType consumes minimal power and requires little board space, making it appropriate for wristwatch-sized devices. But because controlled tilting of one's forearm is fatiguing, a wristwatch using this technique must be easily removable from its wriststrap. A ...

19 What do you have in your walls? 77%

 Alex Perry
Linux Journal October 2002
Volume 2002 Issue 102

The physics, hardware and software behind an easy-to-build probeyou can run with your sound card.

20 Technical columns: ACM SIGACT news distributed computing 77%

 column 5
Sergio Rajsbaum
ACM SIGACT News December 2001
Volume 32 Issue 4

The Distributed Computing Column covers the theory of systems that are composed of a number of interacting computing elements. These include problems of communication and networking, databases, distributed shared memory, multiprocessor architectures, operating systems, verification internet, and the web. This issue consists of four parts:• a survey of SIROCCO'01 by Pierre Fraigniaud,• a survey of POMC'01 by Rui Fan,• a survey of PODC'01 by myself,• the paper

"Paxos Made Simple ...

Results 1 - 20 of 134 short listing


Prev
Page

1

2

3

4

5

6

7


Next
Page

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2002 ACM, Inc.

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

Tables of Contents

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

Search

- ☐ By Author
- ☐ Basic
- ☐ Advanced

Member Services

- ☐ Join IEEE
- ☐ Establish IEEE Web Account

[SEARCH RESULTS](#) [\[PDF Full-Text \(784 KB\)\]](#)[DOWNLOAD CITATION](#)

The design of COMPASS: an execution driven simulator for commercial applications running on shared memory multiprocessors

- Nanda, A.K. Hu, Y. Ohara, M. Benveniste, C.D. Giampapa, M.E. Michael, M.

IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, USA

This paper appears in: Parallel Processing Symposium, 1998.

IPPS/SPDP 1998. Proceedings of the First Merged International ... and Symposium on Parallel and Distributed Processing 1998

On page(s): 503 - 509

30 March-3 April 1998

Orlando, FL, USA

1998

ISBN: 0-8186-8404-6

IEEE Catalog Number: 98TB100227

Number of Pages: xxv+809

References Cited: 14

INSPEC Accession Number: 5907557

Abstract:

Although shared memory multiprocessors are becoming increasingly popular in the commercial market place, the applications used to evaluate such systems in both academia and industry are still predominantly technical applications such as the Stanford SPLASH2 benchmarks. The difficulty in using commercial parallel shared memory applications such as transaction processing, decision support and web server applications has been in simulating the operating systems functions that are heavily used by these applications. We describe the design of an execution driven simulation tool called COMPASS (COMmercial PARallel Shared memory Simulator). We have used COMPASS at IBM to study the behavior of decision support applications and are currently studying the behavior of transaction processing applications and web servers.

Index Terms:

virtual machines shared memory systems parallel processing operating systems (computers) software performance evaluation scheduling COMPASS execution driven simulator commercial applications shared memory multiprocessors academia industry technical applications Stanford SPLASH2 benchmarks parallel shared memory applications transaction processing decision support systems web server operating systems execution driven simulation tool Commercial Parallel Shared Memory Simulator IBM

Documents that cite this document

Select link to view other documents in the database that cite this one.

[\[PDF Full-Text \(784 KB\)\]](#)

[SEARCH RESULTS](#) [\[PDF Full-Text \(784 KB\)\]](#)[DOWNLOAD CITATION](#)

[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) |
[Advanced Search](#)
[Join IEEE](#) | [Web Account](#) | [New this week](#) | [OPAC Linking Information](#) | [Your Feedback](#) | [Technical](#)
[Support](#) | [Email Alerting](#)
[No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to Top](#)

Copyright © 2002 IEEE — All rights reserved

**IEEE Xplore™**
RELEASE 1.4Welcome
United States Patent and Trademark Office[Help](#) [FAQ](#) [Terms](#) [IEEE](#) [Quick Links](#)» [Search Result](#)[Peer Review](#)

Welcome to IEEE Xplore™

- ☐ Home
- ☐ What Can I Access?
- ☐ Log-out

Tables of Contents

- ☐ Journals & Magazines
- ☐ Conference Proceedings
- ☐ Standards

Search

- ☐ By Author
- ☐ Basic
- ☐ Advanced

Member Services

- ☐ Join IEEE
- ☐ Establish IEEE Web Account

Print Format

Your search matched **1** of **807900** documents.Results are shown **15** to a page, sorted by **publication year** in **descending** order.

You may refine your search by editing the current search expression or entering a new one the text box.

Then click **Search Again**.

web <and> compass

[Search Again](#)**Results:**Journal or Magazine = **JNL** Conference = **CNF** Standard = **STD****1 The design of COMPASS: an execution driven simulator for commercial applications running on shared memory multiprocessors***Nanda, A.K.; Hu, Y.; Ohara, M.; Benveniste, C.D.; Giampapa, M.E.; Michael, M.*

Parallel Processing Symposium, 1998. IPPS/SPDP 1998. Proceedings of the First Merged International ... and Symposium on Parallel and Distributed Processing 1998 , 1998

Page(s): 503 -509

[\[Abstract\]](#) [\[PDF Full-Text \(784 KB\)\]](#) **CNF**

[Home](#) | [Log-out](#) | [Journals](#) | [Conference Proceedings](#) | [Standards](#) | [Search by Author](#) | [Basic Search](#) | [Advanced Search](#)
[Join IEEE](#) | [Web Account](#) | [New this week](#) | [OPAC Linking Information](#) | [Your Feedback](#) | [Technical Support](#) | [Email Alerting](#)
[No Robots Please](#) | [Release Notes](#) | [IEEE Online Publications](#) | [Help](#) | [FAQ](#) | [Terms](#) | [Back to Top](#)

Copyright © 2002 IEEE — All rights reserved

"Web Compass" Database Conversion to Web Page of Links on One Topic>Sorted by Title

by Steve Osborne
School of Library and Information Science
Indiana University-Purdue University at Indianapolis

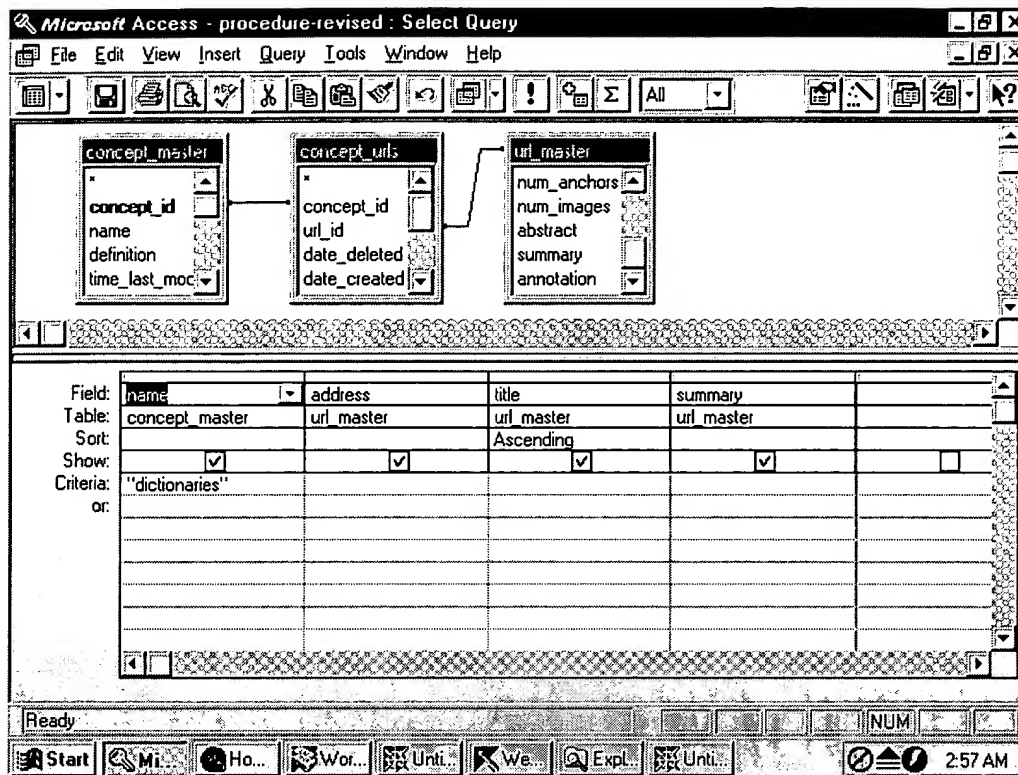
Note: Minor changes prior to any major revision will be in red with a date.

1. Microsoft Access-Setup

- Create a subdirectory in Access or another location of your choosing to store your files.
 - Copy the file:"wc20.wc2" in your WebCompass Directory to the directory you just created.
 - Rename this file with the ".mdb" extension, because if you later copy a newer version the WebCompass database your queries and reports will be lost. To use your old queries when updating your database copy the new one to your working directory then choose 'File'>'Get External data'>'Import'> select your old database>select all your queries and reports (they stay selected when switching tabs) choose Import and you can then use your old queries and reports.10/20/97 10:17:29 PM
 - Goto >File>Open>In "Files of Type" choose 'All Files'>select the WebCompass database
 - The 'Tables' tab should be selected with the six tables used by WebCompass shown
 - Do some exploration on the tables by selecting the table and then either 'Open' or 'Design' on the right side to get a feel for the database
-

• Creating Your Query to extract the URL, Title, and Summary information

- Select the Query tab
- Select 'New'
- Press 'OK' for 'Design View' in dialogue box
- highlight "concept_master" then click 'Add' >repeat for "concept_url" and "url_master"
- close box
- You will see connecting lines between each of the tables meaning that they are related to each other in that an identical field exists in the corresponding table: Note that "concept_urls" purpose is to establish a link between the two tables that contain the fields that we want to use.
- Double-Click "name" field in 'concept_master'; repeat for "address", "title" and "summary" in 'url_master. Note: "name" is the field that contains your search topic.
- Click in 'criteria' row in the "name" column
- Type in your topic exactly as you entered it for your search>press 'enter'
- Click in 'sort' row for "title" column> select 'ascending' in the scroll box in that cell block.
- Your Query should be like below:



- Goto 'File' >'Save as' >type in name for your query
- Close Query Design window
- Highlight your query and chose 'Open' and examine the results

Creating a Report to generate a Web Page of links on your topic

- Preliminary Considerations:
- We are going to generate a HTML report which is different from a normal database report.
- The HTML report will be saved as a 'text' file with the .HTML extension or .HTM if your version of Access does not support long-filenames.
- This HTML file can then be directly viewed in a Web Browser.
- Spacing is not important in HTML but order is: For example,

```
" <HTML>
<HEAD>
<TITLE>Untitled</TITLE>
</HEAD>
```

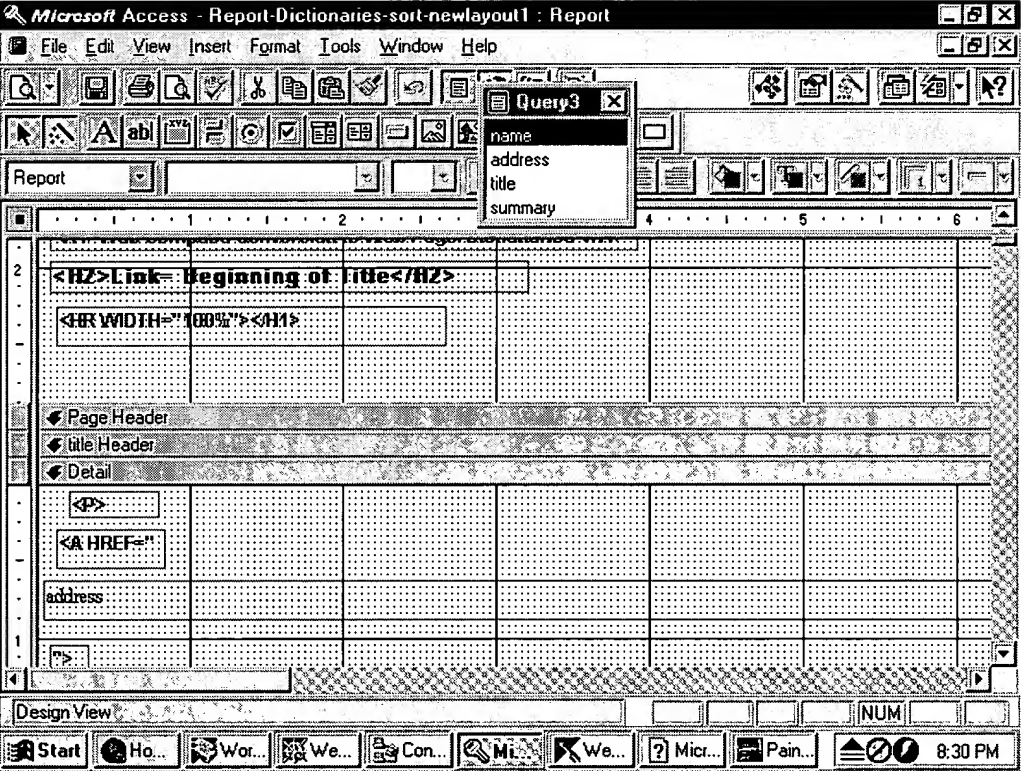
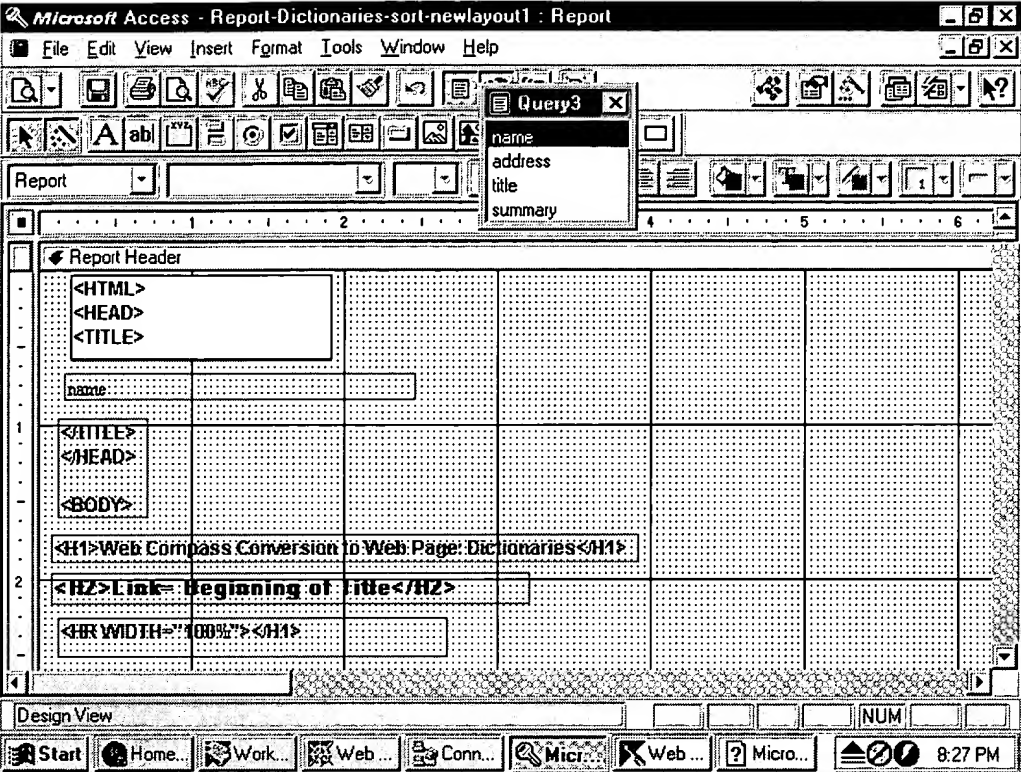
```
<BODY>
```

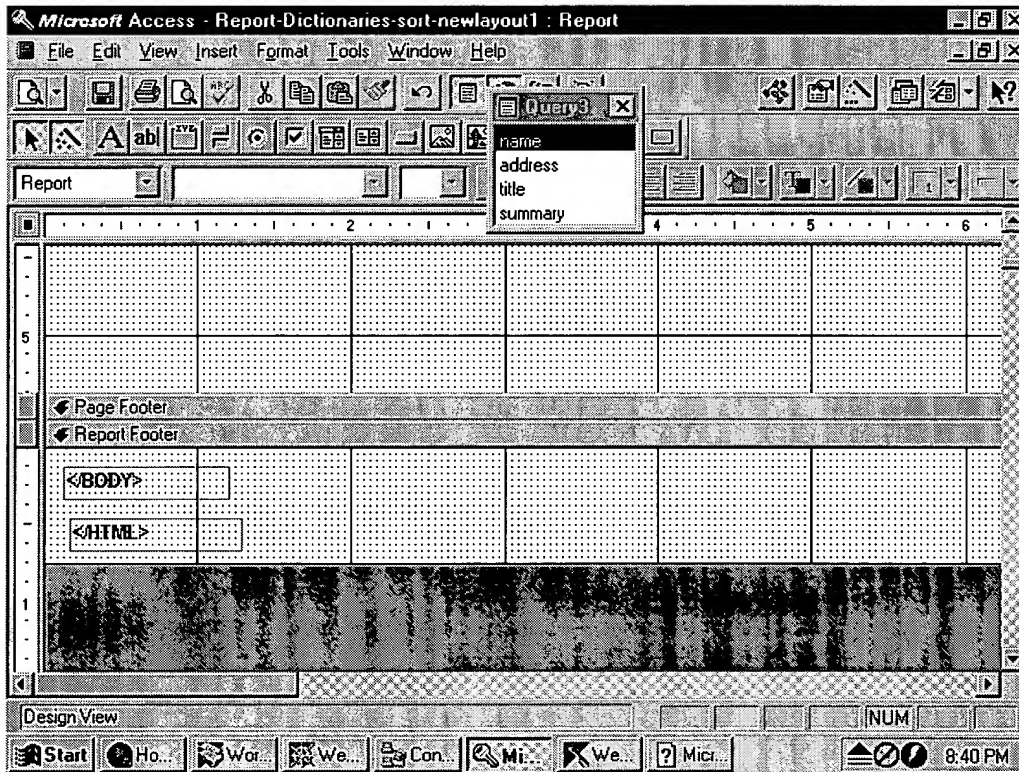
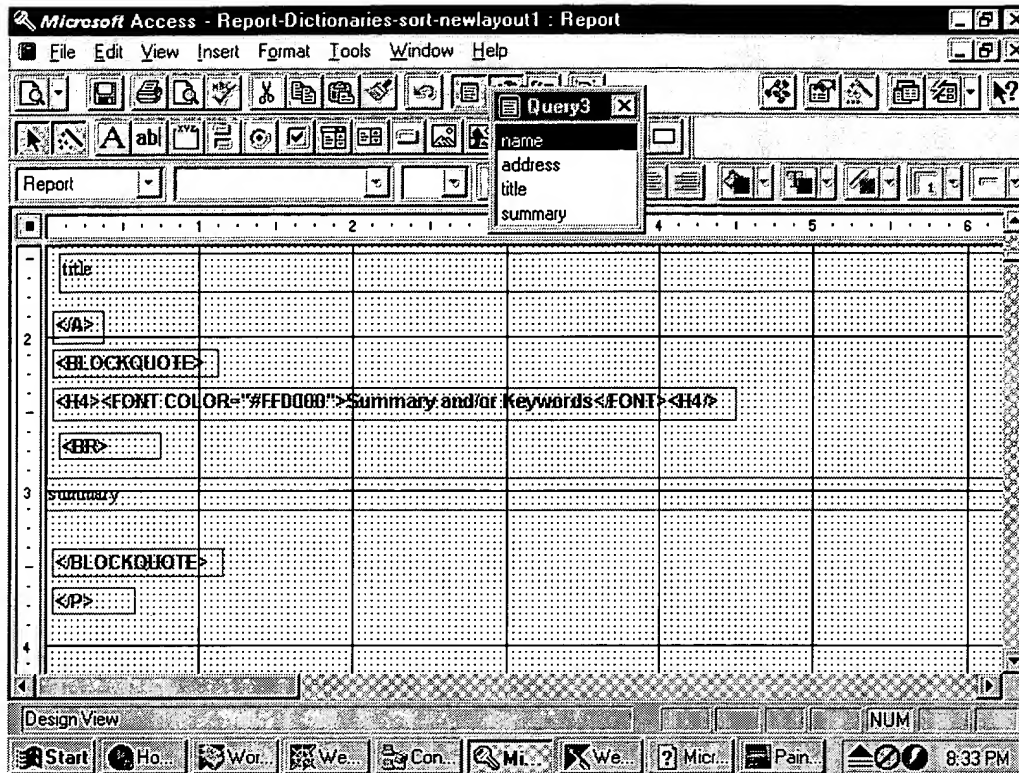
```
</BODY>
```

```
</HTML>" is the same as:
```


"<HTML><HEAD> <TITLE>Untitled</TITLE></HEAD><BODY></BODY></HTML>"

- The fields of the report must be long enough to accomodate the largest size, except for the "summary" field which is a memo field but is defined to extend to the end of the page to save space.
- Select the tab 'Reports'.
- Select 'New'.
- From the "Choose table or query..." box type your query name or choose from the scroll box.
- Highlight 'Design View' and click ok.
- In the 'View' menu select 'Report Header/footer'
- To resize a section of the report go to left scroll bar until a muti-directional arrows appears and drag in the desired direction: Goto the 'Page Footer' section and remove the area assigned it until the 'Page Footer' and 'Report Footer' are on top of each other>This technique is essential so learn it now.
- On the top tool bar, locate the 'hammer and wrench' button , this should be activated. It will appear depressed with a white background. A new toolbar will be activated.
- Going through each step would be too time consuming so you will learn two more techniques along with resizing the sections you will be able to do the rest by following the image of the completed report.
- To enter text (in this case HTML), click on the Large A in your new toolbar. Place the crosshair cursor in the area you want and draw a box the approximate size that you want> it will expand by itself if needed when you type in HTML. Type in your text.
- You can move this box later by clicking and holding down the mouse button> a hand will appear then move where desired.
- to change the size of a box click to activate then move to the little squares around your text box and drag in desired direction.
- To create a field box: in this case "name", "address", "title", and "summary" Goto>the 'Field Button' just left of the 'Toolbar Button' you used earlier> with your left mouse button drag and drop the field into the desired position: put it about 1" away from the left margin to make the part easier.
- The left-hand box with the field name is a label which we are not using this time so select the box and delete it by going to 'Edit' then delete or press the delete key.
- Now expand the left margin by selecting the current edge until the multi-directional arrows appears and drag right to 8.5".
- Resize this field and the other three from the left margin to the right margin with leaving a little space on either side for clarity in editing.
- Follow the next four images from my computer screen or the following text layout for the rest of the report creation:





- Please note there may be a little overlap in the images>use the rulers.
- Text description:
- (Report Header Section)
-

```
<HTML>
<HEAD>
<TITLE>
```

"Name" field box here

</TITLE>

</HEAD>

<BODY>

<H1>Web Compass Conversion to Web Page: Dictionaries</H1>

<H2>Link= Beginning of Title</H2>

<HR WIDTH="100%"></H1>

(Detail Section)

<P>

<A HREF="

"address" field box here

">

"title" field box here

<BLOCKQUOTE>

(next two lines in one box)

<H4>Summary and/or keywords<H4/>

**
**

"summary" field box here

</BLOCKQUOTE>

</P>

(Report Footer section)

</BODY>

</HTML>

**You are now ready to save your work, actually since this procedure is rather lengthy you should save your work frequently. Save your work by going to 'File'>'Save as/Export'>in "within current database" type in new name or click 'ok' if you have already saved. Choose 'Save as/Export' again> choose 'To an external File or Database ' >click 'ok'> in 'Save as type' choose 'text'> type in your filename with the ".HTML" extension> click 'Export'.
Your file will print out for awhile, my current "Dictionaries" is over 200 pages long.**

Open up your Web Browser and in the 'File' menu choose 'open file' or 'open page' and view your Web page.

Good Luck!

If you have any problems e-mail and I'll try to help.

E-Mail to Author

To Examine my generated Web Compass page on Dictionaries:**Web Compass Conversion to Web page of links on a search topic** ©Steve Osborne10/20/97



ARTICLES BYTEMARKS FACTS HOTBYTES VPR TALK

BYTE .com

Search BYTE.com

Search



Navigating with a Web Compass

March 1996 / [Reviews](#) / Navigating with a Web Compass

Categories

Resources

BYTE Humor

Print Archives

About Us

Free E-mail Newsletter
from BYTE.com

your email here



Subscribe

Quarterdeck harnesses leading-edge "metasearch technology to create a smart agent that searches the Web and organizes the results

Rex Baldazo

Searching for a specific subject or topic on the Internet's World Wide Web is like driving cross-country without a road map: You'd better enjoy the journey because you may never reach your destination. Index sites such as Lycos (<http://www.lycos.com>) and OpenText (<http://www.opentext.com>) are a good starting point. But no one index can cover everything, nor can any search site always be up to date; you often end up searching several of these services in t Quarterdeck's new Windows utility called WebCompass greatly simplifies th of searching multiple indices.

Developed by Limbex Corp., WebCompass is actually not a search tool but a *metasearch* tool (see the sidebar "*Searching from Among Searchers*"). That's fancy way of saying that WebCompass does not perform any searches itself. Rather, it runs other search resources such as Lycos and Yahoo (<http://www.yahoo.com>) and the BYTE Site (www.byte.com). When the editor of BYTE magazine saw WebCompass demonstrated at Comdex/Fall '95, we were impressed we gave it our Best of Show award. And now that we've had a chance to give it an extended test drive, we are more impressed than ever.

Use Your Own Browser If You Want

The engine that runs WebCompass is a Common Gateway Interface (CGI) application, which means you must have a Web server running on your local machine (WebCompass includes a version of Quarterdeck's WebServer). You use the Web browser of your choice to control WebCompass ([see the screen](#)). Though Quarterdeck's QMosaic is packaged with WebCompass, configuration problems with our Microsoft Network Internet connection prevented its use. We tried the Microsoft Internet Explorer 2.0 as our interface, which proves just how remarkably flexible WebCompass is: It doesn't care which browser you run.

WebCompass also comes with a Microsoft Access 2.0-format database--as well as 32-bit ODBC drivers--to store and index search results. When you install

WebCompass, you have the option of installing a database that has been prepopulated with various topics to aid in organizing and cross-referencing data.

Add up these components and WebCompass has an imposing footprint. Quarterdeck recommends around 30 MB of disk space for a typical installation, about half of which is the prepopulated database. Once you start storing and indexing your search results, the database will only get bigger.

Resource Management

The key to WebCompass is its ability to interact with various search engines, it terms *resources*. The package comes with several resources configured, including CNN, CNN Sports, Yahoo, Lycos, and Excite. In turn, each resource is grouped into a category, such as General Resources or Technical Resources.

When you run a WebCompass search, you specify both the search terms and resources you wish to use. WebCompass then goes out across the Web and runs those resources for you. Instead of you having to visit each search site in turn, WebCompass does the dirty work of querying the resources, collecting all the responses, and presenting them on a single page ([see the screen](#)).

You can use resource categories to group similar resources. For example, if you're running a query on *Montana* but you want the quarterback and not the state, you can tell WebCompass to search only the Sports resources. If you want to search the resources for Montana, you can do that as well. The search results would include references to Joe Montana, the Big Sky conference, and perhaps Yellowstone National Park.

If the configured resources don't fit your bill, you can add new resources and reassign existing ones to different categories. Adding a new resource is a two-step process. You first enter the URL along with descriptive information and the default category. Then, WebCompass reads in the search resource at that URL. You must then specify which field on the resource page will be used to enter search text. You also have to set up the resource page by entering search criteria.

You set up the search page the way you would want to use it if you were running it directly. From then on, whenever you choose a WebCompass category that includes this new search resource, WebCompass can automatically run that search page for you.

Undercover Agent

WebCompass has two search modes: interactive and agent-based. Interactive is the traditional mode: You type in your search terms, select a resource category, and activate the search. WebCompass then does the hunting and gathering and presents all the results, or *hits*, from the various resources on a single page.

You can investigate each hit in turn and add it to your local database by checking the box next to it. You then can add the hit to a topic that is usually related to the search term you used to find the hit. The WebCompass Agent will then retrieve

index the document in your local database so you have it for future reference. In fact, this local topic database is one of WebCompass' search categories, so you search it just as you would any other resource category.

The Agent is one of the unique strengths of WebCompass. It runs as a separate application outside the browser, though you activate it from the main WebCompass page. The Agent has two roles: It retrieves and indexes articles you have selected as a result of a previous search, and it automatically performs searches for you.

Once you've specified which articles to add to the local database, the Agent can fetch and retrieve them, index them, and organize the documents for later review. The indexing scheme is fairly reliable, usually getting related articles together. In tests using WebCompass, the Agent sometimes put unrelated articles together, but the software generally took a reasonable first stab at grouping the articles into topics.

When you run a search interactively, you can add it to a new or existing topic. For example, I ran a search on the word *Bosnia* and created a new topic also called Bosnia. (Topics are keywords that help organize data in the local database.) If you specify a particular topic as active, the WebCompass Agent will automatically run the searches related to that topic. The software then adds these new search results to the local database for later review. Agent searches can run while you use your browser to view something else.

A Few Nits Found

Having the Agent repeatedly run the same query every day quickly adds up to a large number of articles indexed in the local database. Thankfully, WebCompass comes with tools for managing the database, but since these are also Web-based, they feel a bit clunky at times. For example, you can delete a group of articles with a single click, but to delete a single article you have to jump to a separate form. And you can move articles from one topic to another, but since the interface is Web-based, drag-and-drop commands are unavailable. You have to step through a Hypertext Markup Language (HTML) form to accomplish the task. You also have to use HTML to configure new resources and set up the Agent.

Furthermore, while being able to use different browsers is a strength of WebCompass, it is also a weakness. The caching schemes of some browsers cause a mismatch between what your browser shows and what WebCompass is actually doing. For instance, the browser might indicate that the Agent is inactive when in fact it is out there busily running an indexing task. You can fix this inconsistency by reloading your browser whenever you suspect it might be out of sync with WebCompass.

What's more annoying is that WebCompass is a single-user application, which means you can't set up an instance of WebCompass on your network and let different people maintain their own topic databases using the common installation.

These are minor flaws, some of which will no doubt be ironed out in the next release. What matters is that WebCompass has quickly made itself an invaluable part of our Internet tool set.

PRODUCT INFORMATION

WebCompass 1.0.....\$100 (anticipated stre

Quarterdeck Corp.
Marina del Rey, CA
Phone: (310) 309-3700 or (800) 683-6696
Fax: (813) 523-2335
Internet:
<http://www.quarterdeck.com>

Circle 1049 on Inquiry Card.

HotBYTEs

- information on products covered or advertised in BYTE

Search from Anywhere

[screen_link \(69 Kbytes\)](#)



Quarterdeck's WebCompass search agent lets you launch a search from any browser (for example, Microsoft Internet Explorer 2.0).

Let the Agent Do the Querying

[screen_link \(29 Kbytes\)](#)



The agent queries Internet search sites and groups the results on a single pag

*Rex Baldazo is a technical editor in BYTE's New Media department. You c
reach him at rbaldazo@bix.com.*



Up Level



Next



Search



Comment



Subscribe

COPYRIGHT 1997 Information Today, Inc.

DESCRIPTORS: Association of Information and Dissemination Centers--
Conferences, meetings, seminars, etc.; Information storage and retrieval
systems--Product development

PRODUCT/INDUSTRY NAMES: 7375000 (Database Providers)

SIC CODES: 7375 Information retrieval services

FILE SEGMENT: TI File 148

1/9/5 (Item 2 from file: 47)
DIALOG(R)File 47:Gale Group Magazine DB(TM)
(c) 2001 The Gale group. All rts. reserv.

04830012 SUPPLIER NUMBER: 19383933 (THIS IS THE FULL TEXT)
Surfing corporate intranets; search tools that control the undertow.
(includes related articles on searching databases via an intranet and
intelligent search agents)
Zorn, Peggy; Emanoil, Mary; Marshall, Lucy; Panek, Mary
Online, v21, n3, p30(16)
May-June, 1997
ISSN: 0146-5422 LANGUAGE: English RECORD TYPE: Fulltext; Abstract
WORD COUNT: 9276 LINE COUNT: 00779

ABSTRACT: The availability of commercial and free intranet search engines
promises to meet the unique search and retrieval needs of intranet users.
Free and commercial search engine offerings are reviewed based on their
technical as well as advanced searching capabilities. Users should choose a
search engine based on the type of documents on their site, the site's
size, the number of web servers, the server platform and available
technical expertise.

TEXT:

Intranets consist of web pages, documents, databases, and other
information that sit on a web server or web servers behind an Internet
firewall.

THE INTRANET EXPLOSION

The growth in popularity of corporate intranets over the past year
has risen to epic proportions. A study by the Forrester Group reveals that
over two-thirds of Fortune 500 companies interviewed already have, or are
seriously considering implementing a corporate intranet as a means for
sharing information across their organizations (1). Exactly what is an
intranet and how does it differ from the Internet?

Intranets are internal corporate networks set up to take advantage of
popular Internet communication protocols such as TCP/IP and HTTP, and other
Internet tools such as web servers, web browsers, and HTML. While the
Internet largely provides public, unrestricted access to its content,
intranets strictly control access to content, allowing authorized users
only. Intranets consist of web pages, documents, databases, and other
information that sit on a web server or web servers behind an Internet
firewall. Employees use a standard browser, the same browser they use to
access information on the Internet, to search and locate internal
information. These web sites are devoted to providing access to internal
information to employees, while keeping their content secure from the rest
of the Internet community.

Intranets are popular with corporations for many reasons:

- * Intranets can be easier and cheaper to implement than traditional

groupware solutions like Lotus Notes. Web server and browser software is inexpensive--in some cases, free--and can easily be loaded and operational on a corporate network running the TCP/IP network protocol within a matter of hours.

- * Intranets are scalable--they can start out small with just a few links or home pages in place, and can then grow easily over time to include a huge variety of information with little or no additional investment in infrastructure.

- * Intranets are built using open standards that allow a variety of PC platforms (Windows for Workgroups, Windows 95, OS/2, Macintosh, UNIX, etc.) to access the same information.

- * Intranets can incorporate access to a variety of document and media types including Adobe Acrobat (PDF) documents, HTML, word processing, spreadsheets, sound, video, and graphics applications.

- * Empowering employees to independently locate and use a variety of information ranging from online phone directories to Human Resources information can save time and foster employee satisfaction--two things of value to most organizations.

- * Finally, intranets allow corporate users to capitalize on their knowledge of using the Internet by using the same software to locate and access internal information. End-users are becoming increasingly comfortable with the web browser and its hypertext linking as an interface to all types of services and information.

SURFING THE INTRANET?

As the content of intranets increases, so does the need for tools that help users locate the information they're looking for quickly and easily. Typical Internet users would find it very difficult, if not impossible, to locate and return to sites they find useful if they didn't use tools like the bookmarking feature of most web browsers and publicly available indexes and search engines such as Alta Vista, Lycos, Open Text, and Yahoo!

The same principles hold true when applied to locating information on intranets. Even the most careful and organized intranet webmaster will find it tough to come up with an organizational scheme for the enterprise web site that makes sense to all users. In addition, the more content added to intranets, the more levels of organizational hierarchy the user will be required to "drill down" through before locating relevant content.

While some users of the Internet may be willing to spend time "surfing" to locate needed information, corporate intranet users and their employers are much more demanding. Companies do not want their employees using unnecessary time to locate needed information in the increasing sea of documents and data available from corporate intranets. By the same token, employees that may be very patient when trying to find information on the giant Internet will not feel that same patience when trying to locate a specific piece of vital internal information.

ENTER...INTRANET SEARCH ENGINES

To help organizations find solutions for locating information on intranets, several intranet search engines--both freeware and commercial--have been developed to address the unique search and retrieval needs of intranet users and developers. These search engines are designed to crawl and index internal web servers and/or portions of these servers to create custom, searchable indexes of the documents and data housed on the servers. They have some features in common with the very large and very popular general Internet search engines, but they also contain some unique capabilities that set them apart from their Internet search engine counterparts.

While both Internet and intranet search engines provide indexing for

basic HTML documents, intranet search engines often also provide indexing for other document formats (PDF, word processing, spreadsheet, graphics, databases, etc.) that may be contained in an intranet web site.

Internet search engines are often measured by their ability to provide access to the largest Web indexes, and retrieval from many Internet searches can be overwhelming. Intranet search engines are usually designed to provide more precise data filtering and retrieval, limiting the amount of information the user is required to sift through. To do this, the actual indexing process of an intranet search engine is probably deeper than its Internet counterpart.

According to a recent article in InfoWorld, "...the companies that can configure their search engines for better relevance in search results will be the winners in the intranet field. That difference will come from how their search engines house information" (2).

Corporate librarians and other information professionals can play an active role in evaluating and recommending the use of such products in their organizations; they need to familiarize themselves with the products available and the issues surrounding their selection, implementation, and use. This article takes an in-depth look at the major free and commercial intranet search engines currently available and analyzes differences in features, cost, ease-of-use, and hardware and software requirements. In addition, we'll take a look at some trends that we see affecting the intranet search engine and web server industry that may influence the availability and functionality of intranet search engines in the future.

Information professionals familiar with the indexing and searching process can lend a lot to the evaluation and implementation of intranet search engines within organizations. In-depth knowledge of searching techniques, including use of controlled vocabulary, Boolean operators, proximity operators, and relevancy ranking, is necessary for evaluating the potential effectiveness of the various intranet search engines available. An understanding of, and experience with standard indexing practices and parameters can also ensure that the data contained in the various indexes built on a corporate intranet will facilitate accurate and efficient data retrieval. As the information in intranets grows, the importance of having powerful, accurate, and comprehensive search tools becomes one of the most important issues facing organizations.

RATING THE SEARCH ENGINES

A wide variety of free and commercial search engines for use with intranets exists; products vary greatly in support for search and retrieval, operating system platforms, web server environments, file formats, and cost. This article provides a detailed analysis of eight search engines that can be used for indexing and retrieving documents and information located on an intranet, including:

- * Alta Vista (both Alta Vista Search INTRANET Private eXtension and Alta Vista Search INTRANET XL Private eXtension)

- * Excite for Web Servers
- * Fulcrum Surfboard
- * Glimpse/Harvest
- * ht://dig
- * Open Text Livelink Search
- * Verity Topic Search
- * ZyIndex Webserver

For each search engine covered, we took an in-depth look at technical functionality, searching features supported, results display, and cost. Technical functionality criteria include:

- * server platforms supported (UNIX, NT, VMS, etc.)
- * scalability (can it index an entire intranet, including multiple

web servers, and then also build specific indexes based on directory, file type, etc.?)

- * indexable file types (HTML, PDF, word processing, spreadsheet, distributed databases, etc.)

- * technical support availability (toll-free 800 number, World Wide Web site, email, support contracts)

- * price/licensing across multiple sites/web servers

Advanced searching and results display options were also looked at closely. Features focused on include:

- * Boolean logic, including nesting

- * proximity and phrase searching

- * truncation

- * search set manipulation

- * duplicate detection

- * field searching

- * thesaurus or concept searching

- * file formats supported in results display: (i.e., can documents be displayed in results display in native format or just HTML?)

- * relevancy ranking or results sorting

- * keyword-in-context (KWIC) display

Detailed descriptions of the features and functionality of each product examined follow. They are divided into two categories: 1) free, and 2) commercial. In addition, a chart comparing the features of all the search engines covered in this article is provided (Figure 1).

(Figure 1 ILLUSTRATION OMITTED)

THE FREE (OR ALMOST FREE) SEARCH ENGINES

Excite for Web Servers (<http://corp.excite.com/ews.html>)

Excite for Web Servers (EWS) is a free product from Architext Software Inc. It is based on the Excite search engine available for public use on the Internet. This standalone product supports a variety of server platforms, including UNIX, Windows NT, and Macintosh. Document collections, specified by the local system administrator, contain the information about what is to be indexed. These document collections are comprised of CollectionContents, which are lists or descriptions of files to be indexed, as well as the CollectionIndex, which is the searchable index of these files. EWS currently only indexes HTML and text files, but there are plans to index PDF files in the future.

The initial release of EWS was somewhat limited in its scalability, but this has been corrected in version 1.1 of the software. It is now possible to index files on multiple servers, but there is a limitation on individual file sizes and some question about performance with very large document collections. EWS says that several customers have collections larger than 1GB, but as collections grow, performance is impacted with slower search speeds. It does not appear that EWS can include files outside the local intranet, given the indexing options available for the CollectionContents.

Although EWS is a free product, purchase of a maintenance contract is necessary for technical support of the product. Cost for support at this writing (December 1996) is \$995 per machine per year. This support includes free upgrades, as well as email and phone support. A spidering version is slated for release in the first quarter of 1997. This version will be provided free to customers with current maintenance contracts.

Searching features supported in EWS are fairly basic (Figure 2). The search interface is customizable, with the main search mode as either concept-based searching or keyword searching. A concept search allows you to enter a phrase such as pro-choice vote in Michigan.

(Figure 2 ILLUSTRATION OMITTED)

Results are then returned with a confidence rating. However, some results may appear not to contain the terms entered in the search because the concept search attempts to identify concepts rather than the exact terms entered.

Using advanced statistical methods EWS analyzes documents for relationships between terms, and uses those relationships to identify search concepts. EWS stresses that concept searching does not use a thesaurus, which it feels limits precision in results returned. The keyword search functions similarly to a Boolean AND search. Traditional Boolean searching is offered with the current release of EWS; truncation is provided through automatic stemming. Search set manipulation and field searching are not supported at this time.

Results are displayed by relevancy with an option for viewing a summary of the document. The summaries give a brief abstract of the document so it is not necessary to access the actual document to determine if it satisfies the query. EWS has a Query by Example (QBE) feature to find more documents like those returned. Other features include subject groupings when broad topics are searched, thereby providing the user with a means to refine the search. There is a facility for duplicate detection.

EWS is currently being used to index several public Web sites including those of Nestle, Adobe Systems, and Bell Industries. The search screen and results display is essentially the same at these sites, with some variation depending on the design of the page. EWS is an excellent choice for indexing small collections of documents on a single web server. Organizations just beginning to implement an intranet or small organizations only wanting to index HTML, text, and PDF documents would find EWS to be a reliable search engine. It seems fairly easy both to implement and to administer, and purchasing the maintenance contract would provide these organizations with added support as their intranet grows.

[ht://Dig](http://Dig) (<http://htdig.sdsu.edu/>)

[ht://Dig](http://Dig), developed at San Diego State University, is a free and complete web indexing and searching system for an intranet. This standalone product can cover several different web servers at a site but is restricted to the UNIX server platform. As long as the web servers understand the HTTP 1.0 protocol, the web server will work with [ht://Dig](http://Dig). There is no mention of moving to other server platforms in the various documentation describing [ht://Dig](http://Dig).

The software can index all HTML and ASCII text files; other file types are supposed to be searchable in future versions. An interesting security feature of this free software is its ability to search a protected server when the correct password is given. Despite a lack of advanced searching functionality, this security feature is a plus for corporate intranet use.

Samples of how the search engine works can be found from the [ht://Dig](http://Dig) home page by linking to San Diego State University's home page (<http://www.sdsu.edu/>). Basic and advanced searching screens are shown, as well as specific indexes that search through specialized subsections of the university. These links to actual working databases provide an excellent understanding of [ht://Dig](http://Dig) searching capabilities and show the scalability of the program. It is possible to set the software up to search an entire intranet, or a smaller subsection. In addition to searching examples provided through San Diego State University, the program is available immediately for downloading from the [ht://Dig](http://Dig) home page (<http://htdig.sdsu.edu/>).

Search capabilities of [ht://Dig](http://Dig) are fairly basic. From the advanced search screen, it is possible to specify "match all" (AND), "match any" (OR), or "match Boolean" (which accepts the terms AND and OR as commands,

plus allows for nesting using parentheses. There are no options for search set manipulation, field searching, or proximity or phrase searching. Truncation is automatic, with no option for searching the root term only.

From the search results screen, your search terms are highlighted. You have the option for long or short results, with the most relevant terms receiving more stars. The current search strategy is listed at the top of the results page for easy reviewing and refining of your search.

Some other search capabilities include the ability to create a controlled vocabulary list by adding keywords to HTML documents, and the ability to do "fuzzy searching," which provides algorithms for search result enhancements, such as finding synonyms.

System requirements and installation notes are clearly listed from the [ht://Dig](http://Dig) home page. Although the notes are quite complete, knowledge of the UNIX operating system and code compiling is necessary. This system, although not appearing extremely difficult to install, is not turnkey. There are files to download, directories to configure, and scripts to modify. There is a large configuration file for customization once the software has been installed.

Technical support consists of a newsgroup of users; it seems helpful, and an archive of the newsgroup messages covers several common problems. Additionally, there is detailed online documentation, and an email address for Andrew Scherpbier, one of the creators of [ht://Dig](http://Dig).

Harvest/GlimpseHTTP/ WebGlimpse (<http://glimpse.cs.arizona.edu/webglimpse/>)

Harvest is a collection of UNIX-based Internet tools designed to perform several different tasks, such as gathering, extracting, and replicating Web information. The project that created Harvest is officially over, with funding that ended August 1996, although parts of the software collection continue as commercial ventures or as supported by volunteers. One part of this software group is a searching facility that can be applied to both Internet and intranet use.

The search engine software is called Glimpse; it is available and supported in the forms of Glimpse, GlimpseHTTP, and WebGlimpse. In order to use Glimpse on a web site (whether internal or external), you need either GlimpseHTTP or WebGlimpse. According to the GlimpseHTTP Web site, though, WebGlimpse does a superior job of browsing and searching on a single web page. Additionally, WebGlimpse has the capability of indexing and searching several Web servers at once, which GlimpseHTTP cannot do. This section of the evaluations will focus on WebGlimpse, since it is the most appealing to those considering an intranet search engine.

When installed, WebGlimpse inserts a search box at the bottom of every HTML page specified. The search box can be set to search the entire index or the "neighborhood" of the page. The "neighborhood" is defined by the installer as a certain number of links away from the current page. Both this box and the advanced searching box supports Boolean (AND, OR, and NOT), but it is command-based rather than form-based. Thus, someone trying to find Web pages containing the phrase "Arizona Desert" and the word "Windsurfing" would have to type in `Arizona desert;windsurfing` as a search command.

The advanced searching page also includes options for case-sensitive searching, partial-word searching, and the ability to match misspelled words. In WebGlimpse, only HTML and text pages can be searched. Harvest has more search format capabilities, but as you move away from WebGlimpse, you also move away from a complete, supported product.

In WebGlimpse, there are no options for field or concept searching, proximity searching, detecting duplicates, or manipulating search results. From the advanced screen, it is possible to specify the maximum number of

files that you would like to have returned by your search.

Your search results include the title of (and a link to) the URL, the date it was last modified, and the list of all lines that matched the query. This results screen thus produces a modified keyword-in-context (KWIC) display, which is extremely useful in determining the relevancy of your retrieval.

Currently, there are no sample databases to search using WebGlimpse. The developers are in the process of releasing a new version, and the "practice" searching was not yet available at this writing. Still, there is a spot on the WebGlimpse home page for sample searching in the future. The entire source code is available for downloading, as is a series of executables for installing the program.

Other Free Intranet Search Utilities: WAIS, htdgrep, and SWISH

Other free utilities and search tools are available for use on intranets. However, most require at least some knowledge of CGI scripting, PERL, and/or another programming language to customize for use at your site. In addition, most were developed for the UNIX platform, using utilities and tools available in the UNIX environment that may or may not be transferable to other platforms. Although many were developed using either the PERL or C programming languages, both of which are available for most platforms, the conversion process to other operating environments can be painful without the right expertise. Often, the complexity of the searching that can be done with these free tools and utilities depends on the programming expertise available at your site.

Probably the most widely recognized and powerful searching utility available is WAIS (Wide Area Information Service). WAIS grew out of a project started by Apple Computers, Thinking Machines, and Dow Jones and became one of the most widely used searching tools in the early days of the Internet. WAIS evolved for use with the Web, and a Web version ([www.wais](http://www.wais.com)) is available. WAIS can support fairly advanced searching features such as Boolean, phrase, field, and proximity searches, as well as truncation. There are many varieties of WAIS now in use in addition to [www.wais](http://www.wais.com), including freeWAIS, Son of Wais, Kid-of-WAIS, and a commercial version available from WAIS, Inc. For more information on WAIS or to download the necessary files to get started, see <http://www.eit.com/software/www.wais/www.wais.html> or http://ls6-www.informatik.uni-dortmund.de/ir/projects/freeWAIS-sf/fwsf_1.html.

Another popular tool used to create searchable indexes on intranets is htdgrep. Htdgrep is a UNIX-based CGI script written in PERL that allows queries to any document accessible to your HTTP or web server on a paragraph-by-paragraph basis. Htdgrep allows users to create forms-based HTML files that pass all search parameters specified to the searching script. Most sites using htdgrep write their own CGI scripts, adapting htdgrep to meet their needs and hard-code searching options such as Boolean, truncation, and case-sensitive searches. A FAQ on htdgrep is located at <http://iamwww.unibe.ch/~scg/Src/Doc/htdgrep.html>.

SWISH (Simple Web Indexing System for Humans) is a C program designed to index directories or individual files (usually in HTML format) and provide a search interface to the index created. SWISH uses a configuration file to specify directories and files to search, stop words, and some other basic parameters. SWISH supports Boolean searching and relevancy ranking of results, but not truncation. As is, SWISH can be executed from a command line interface. To use SWISH with an HTML forms interface, you will need to write a CGI program that acts as a gateway between the SWISH program and passes it the necessary searching parameters. To learn more about SWISH or to download the source code, see <http://www.eit.com/software/swish/swish.html>.

These utilities for creating intranet search engines are only the tip of the iceberg in terms of what is available out on the Internet for creating and customizing search engines for use on intranets. There are many more ways to implement intranet search engines, depending on your needs and willingness to program and customize.

COMMERCIAL INTRANET SEARCH ENGINES

Alta Vista (Alta Vista Search INTRANET Private eXtension and Alta Vista Search INTRANET XL Private eXtension) (<http://altavista.software.digital.com/>)

Alta Vista has developed an intranet search engine that uses the same technology that powers the Alta Vista Search Public Service, the popular Internet search engine developed by Digital Equipment Corporation and released in December 1995. Alta Vista Search INTRANET Private eXtension is available in two versions, PX and XL PX. The PX version is for smaller machines; pricing starts at \$16,000. XL PX is intended for machines with 2GB or more of memory; pricing starts at \$66,000. The software is currently available for Alpha UNIX or Digital Alpha servers, but a version for Windows NT is expected soon.

Like the public search engine, Alta Vista Search Intranet Private eXtension indexes every word on all the web servers on the intranet, as well as specified Internet sites. Using spider software, the search engine crawls the servers behind the company firewall as well as any external Web sites specified, creating an index of every word. This feature is particularly useful to libraries wanting to provide access to information on selected public Web sites through the intranet search engine. Because the software supports multinational intranets, it is able to index servers in multiple locations.

Alta Vista plans to support a wide variety of formats, but the current release indexes only HTML and text files. Database indexing is available with an add-on product--Alta Vista Search Toolkit. Setup and maintenance of the software is designed to need little administration.

Search features and the search interface are the same as with Alta Vista's Internet search engine: Boolean, proximity and phrase searching, field searching, and search set manipulation are available from a forms-based search screen. The results are displayed in relevancy order and can be displayed in standard, detailed, or compact format.

Because the product is relatively new (November 1996 release), it is hard to know how well it will be received by the corporate community. Given the excellent search features offered on the publicly available search engine, as well as its popularity, it is likely to attract a great deal of deserved attention.

Fulcrum Surfboard (<http://www.fultech.com/>)

Fulcrum Surfboard is available from Fulcrum Technologies of Ottawa, Canada. Surfboard is an add-on to SearchServer, Fulcrum's multiplatform search engine driving several Fulcrum products, including SearchBuilder and Find. Supported server platforms are Windows NT and UNIX with support for most any CGI compatible web server including those available from Netscape and Microsoft. The software incorporates security features that limit access based on current firewall specifications or other security needs. It also has a reporting feature so that vital information can be gathered about employee intranet use.

Suriboard has a distributed search architecture that supports open system standards including the use of Z39.50 standards in its search protocol. Fulcrum intentionally built its product using industry standards so that it is able to operate with a wide range of system components. The index is maintained on web servers using MultiGate, Fulcrum's gateway application that accepts search requests, queries the Surfboard index, and

returns the results as HTML documents to the end-user. MultiGate is able to access both local and remote sites within the corporate intranet, as well as public Web sites.

Installation is designed to be simple with GUI-based administration tools and wizards to guide the system administrator through setup. However, knowledge of SQL is necessary for maintenance and support. Cost is \$6,250 per server for Surfboard plus an additional \$5,000 for SearchServer and \$295 per seat. Technical support is available through a purchased support contract. This contract includes email and phone support as well as administration courses.

Surfboard indexes and searches most document formats, including HTML, PDF, MS Office documents, relational and WAIS databases, NetNews, and over 50 other document formats. The index which Fulcrum creates is actually a series of tables that hold document attributes as well as a pointer to the document. The documents themselves remain in their original location. When a search is submitted, SearchServer queries the tables and returns a list of documents with links indicating their original location.

Advanced searching features include Boolean, phrase searching, truncation, date range searching, structured (field) searching and common language searching (Figure 3). Common language searching allows the user to enter a question: How do I install Fulcrum Surfboard? instead of formulating a Boolean search: install* and fulcrum and surfboard. Surfboard offers a feature called SearchObjects for users to bookmark queries or for administrators to design queries for easy access to commonly requested documents.

The search interface in the demo available from the Fulcrum home page is basic, but design of this interface is customizable as is the results display. Results are displayed in a relevancy ranking with search terms highlighted. Documents are converted to HTML on-the-fly if the original format software is not able to be launched from the Web browser.

Customers of Fulcrum include major players in the information technology field such as Microsoft, CompuServe, and Netscape, as well as a variety of other clients. Because Fulcrum offers several information retrieval products and incorporates industry standards into these products, it is attractive to administrators of corporate information tools, particularly corporate intranets. Its distributed architecture, scalability, and ability to be customized make it an excellent choice for organizations with large document collections in a variety of formats and locations.

Open Text: Livelink (<http://www.opentext.com/>)

Open Text Corporation, located in Waterloo, Ontario, Canada, has developed an intranet suite of products called Livelink Intranet. Livelink Search is the search engine of Livelink Intranet; the other three components that complete the Livelink Intranet family include Livelink Library, Workflow, and Collaboration. Many people may already be familiar with Open Text's presence on the Internet via their Open Text Index on the Web (<http://index.opentext.net/>). Livelink Search uses the same full-text indexing software it uses to search the Internet and includes an option to provide both intranet and Internet searching from its default intranet search screen.

Livelink Search currently supports the following server platforms: Windows NT, SUN Solaris and Sun OS, HP/UX, AIX, SGI, and DEC OSF1. The search engine is scalable and guarantees support of document collections of any size. According to a recent Canadian Newswire release, the new search engine is built to handle tens of gigabytes of information, as opposed to hundreds of megabytes common with other search engines (3).

Livelink Spider is the crawler software that locates the documents on

the corporate intranet and external Web sites and locally indexes their full text. Documents from relational databases, flat files, HTML, SGML, and 40 other common office data formats can be indexed. It also has the capacity to index Internet mail files and Internet newsgroups. Livelink Spider can be configured to "crawl" to specific domains, server directories, and file types, and conversely, it can be configured not to crawl to specific domains or server directories. Livelink Search has the flexibility to support multiple indexes on one server and multiple indexes on multiple servers, making decentralized collections of information easily searchable.

Open Text offers a variety of support for their products, including training courses for end-users, administrators, and developers; and online reference information and user guides. Customer Service Representatives are available to support any questions regarding functionality, use, and configuration of Open Text products; however, in order to take advantage of this service, you must subscribe separately to their Customer Assistance Program. At the time of this writing, the price for Livelink Search and Livelink Spider is \$12,000 and \$12,500 per server, respectively. Netscape's Commerce Server communications software is also included in the package.

Full Boolean searching (AND, OR, and NOT) is supported by Livelink Search, as is proximity searching (NEAR), advanced similarity searching (find more results like this one), truncation with a wildcard (*), and full-phrase searching (phrases with no stopwords). Keyword searches look for literal matches of the words, and concept searches use a thesaurus to locate related terms. Searches can also be run to query a specific field of a document (Figure 4). If a Search Application Programmable Interface (API) is purchased and developed for Livelink Search, results can be manipulated for further advanced searching options.

Retrieved results are ranked based on an intelligent ranking algorithm and can be viewed in three different formats: simple ASCII, on-the-fly HTML, or in native format. Livelink Search can convert non-HTML documents on-the-fly so that they can be viewed by any web browser. Document summaries, if not originally provided, can be created using an automatic document summary generator. There is also an option to view the keywords in KWIC mode, where the keywords are highlighted and the user can easily see where the keyword(s) occurs in the retrieved document.

In addition, searches can be restricted to query-specific sections of documents, because the search tool does index documents based on tags or database fields. The client interface is customizable to suit the needs of the users' searching preferences.

Verity SEARCH'97 (<http://www.verity.com>)

Verity, founded in 1988, is the developer of the Topic family of search and retrieval tools for the enterprise and the Internet. In the Fall of 1996, Verity relaunched its entire suite of search and retrieval tools under a new name: SEARCH'97. SEARCH'97 is a comprehensive, flexible platform for deploying search applications across the corporation. The Verity indexing format is being used by over 500 companies worldwide. Some of the companies bundling the search engine into their software include: SAP, Lotus Notes, Individual Inc., Adobe Acrobat, Documentum Inc., Xyvision, Netscape servers, Dow Jones, Reuters, and Ziff Davis

SEARCH'97 can index information from virtually any document format that has been used in the last ten years, including relational databases like Informix, Sybase, and ODI. It is also working towards indexing data from data management files such as Lotus Notes, SAP, Informix, and Documentum. Touted as the mechanism to harness the "corporate memory" of an enterprise, SEARCH'97 facilitates the collection, management, and retrieval of information throughout a corporation and specified sites on the

Internet, and makes the data available at an employee's desktop.

The SEARCH'97 platform includes a variety of components: SEARCH '97 Personal, Information Server, Agent Server, Advanced Search and Query enhancements, and Knowledge application tools and advanced navigation. SEARCH'97 Personal is an interface used to initiate search queries, access search agents, and implement searches. SEARCH'97 Personal can locally index Internet Web sites at the individual's computer so that personal Internet Web sites can also be queried along with remote corporate indexes. It is supported from within a web browser or Microsoft Exchange. Results can be viewed in virtually any file type, even if the native application is not available locally. SEARCH'97 Personal is available for UNIX, Mac, Windows 95 and NT.

At the center of the SEARCH'97 framework is the Information Server. The Information Server indexes and manages corporate information--the "corporate memory"--and uses a web browser or SEARCH'97 Personal as an interface. A web spider is also included to add corporate data and/or Internet sites automatically to the main index.

The full text of documents is indexed; the indexes are updated automatically when data is added, changed, or deleted. The indexing tools support access to virtually any document format including common office document formats, HTML, PDF, and ASCII text. Remote indexing is available to store information from different sites throughout a corporate intranet. The Information Server also acts as the integration point for advanced searching components such as the agent server, enhanced query, visualization, and knowledge and navigation tools. According to Verity's Product Brief, the following platforms support Information Server: Solaris, IBM AIX, HP/UX, Windows NT, DEC Win Alpha, and DEC UNIX.

SEARCH'97 Agent Server automates the search and retrieval process for the individual or corporation. A search profile is prepared and the agent notifies the requester when information or data match the search profile. Individuals customize their information profiles with a set of keywords, specific sources (Internet or intranet sites, including databases) which the query will be run against, and the preferred method for notification (via email, web page, or pager). The agents run continuously and instantaneously alert the user when new information has been added to any of the sources specified in the search profile. Hundreds of thousands of agent profiles can be initiated per server. SEARCH'97 Agent Server currently operates on Solaris 2.5 and Windows NT 3.5.1 platforms. General availability for Agent Server is scheduled for first quarter of 1997 with pricing at approximately \$70,000.

A Technical Support site ([http:// www.verity.com/tech-support/index.html](http://www.verity.com/tech-support/index.html)) is available from the Verity home page. This site includes a technical support information sheet providing phone and fax numbers as well as email addresses for Verity offices worldwide. The technical support information sheet also details the procedures for obtaining technical support. Online information is searchable and includes FAQs, selected data from their Help Desk database, and selected technical notes. Verity also offers a number of educational courses (<http://www.verity.com/educ/index.html>) for their products. Courses are taught at Verity Training Centers in Sunnyvale, CA and Fairfax, VA or can be conducted on-site.

The Verity search engine offers both literal text and Boolean searching capabilities. Other searching options are customizable using standard web forms.

For literal text queries, commas placed between key terms will search on any of those keywords (implied OR). Truncation of words occurs automatically; however, a specific word or phrase can be searched simply by

placing quotation marks around the word or phrase. A wildcard can also be used to find variant letters at the beginning of a word or letters. Field searching is available for querying a specific date or author. Proximity operators are also supported; search terms can be specified to show up "near" each other, in the same phrase, sentence, or paragraph. A thesaurus is available to retrieve synonyms for additional search terms.

Natural language queries and query by example (find me more like...) are also supported. The search engine takes the user's query, whether literal or Boolean, and supplements it with "fuzzy logic"--an operator that calculates a "more the better" score to determine relevancy ranking.

Multiple collections (or indexes) can be searched simultaneously by the individual and are selected from a list or drop-down menu. The user can also determine the number of results returned from the search query.

Documents returned are given a score and listed in order of relevance. Results can be previewed using a rich-text translator, or displayed in the "native format" that can range from data in an Oracle database to Lotus Notes documents to Adobe Acrobat PDF files. Native formats can only be viewed when the requested file format is available locally on the user's computer, or when a suitable viewer is used.

Additional add-on components to the basic SEARCH'97 can increase the flexibility of searching and improve the relevance of results. These optional intelligent search components include: Enhanced Query, Visualization, and Knowledge and Navigation Tools. Enhanced Query uses query technologies such as natural language processing (NLP) and query by example (QBE). A user can type in a search in the form of a question and then use NLP to locate information based on the phrases that were entered. In using QBE, a searcher can copy an example of relevant text from a retrieved result and paste that text in the search form. The QBE engine will then reformulate the search and locate information relevant to the text that was submitted.

The Visualization components (clustering and summarization) make it easier for users to identify relevant information. Clustering organizes the retrieved results into groups based on commonality of terms. The Summarization component creates an overview of individual documents based on an algorithm that determines the significance of the sentences that make up the documents. These summaries are more sophisticated than the typical summaries created by just the document title and following few lines of text.

Navigation tools allow the user to move through documents more easily by using hyperlinks from one document to another. To further facilitate knowledge transfer within an organization, a systems administrator can use Verity's Knowledge Tools. These tools allow administrators to create their own knowledge bases specific to their business environment that include, but also extend beyond the typical functionality of dictionaries and thesauri. These navigation tools would provide more precise search results by filtering out and eliminating irrelevant documents.

ZyIndex Webserver (<http://www.zylab.com/>, <http://www.zylab.nl/>)

ZyLab International, Inc. was founded in 1983 with the introduction of PC-based full-text indexing and retrieval software. Today, ZyLab offers complete web-based publishing and indexing solutions in its ZyIndex Webserver and ZyImage Webserver product lines. ZyIndex Webserver, the software package we will be focusing our attention on in this section, provides full-text indexing of document collections in over 30 formats and makes them searchable through the Internet or corporate intranet. ZyImage Webserver, a companion product to ZyIndex Webserver, combines the ZyImage scanning interface for OCR (optical character recognition) of documents in electronic format with the powerful indexing and search and retrieval

engine of ZyIndex Webserver.

All technical specifications and searching functionality described apply to both ZyIndex and ZyImage Webserver. The main difference between the two products is that the ZyImage Webserver offers users the additional benefit of being able to view images of scanned documents with an easy-to-use scanning interface. ZyIndex Webserver sells for \$5,995 complete, while ZyImage Webserver, which includes the ZyImage OCR and ZyIndex software, sells for \$11,200 (price includes annual update service). Both products are licensed to cover a total intranet site. ZyLabs offers a full range of technical support options including an 800 number, electronic mail, Web site, and support contracts.

ZyIndex Webserver supports the most popular HTTP servers, i.e., Microsoft and Netscape. However, ZyIndex Webserver can be used with any existing web server product that is HTTP 1.0 compliant, running on the Windows NT platform. ZyIndex Webserver provides a proprietary API that handles the search and retrieval process and interfaces with the document index created during the configuration process. In addition, ZyIndex Webserver comes with a set of HTML templates designed to function as the search forms used by end-users through their web browsers, and as the default display format for return and viewing of search results. These templates can be customized to meet client needs.

The ZyIndex Webserver allows clients a great deal of flexibility in indexing features and document and index security. ZyIndex can index document collections located anywhere on the corporate network and can support more than 30 native file formats including all major word processing programs (Word, WordPerfect, etc.), group 4 TIFF, popular database formats (dBase 3 and 4, FoxPro, etc), Lotus, Excel, EPS (encapsulated Postscript), and ASCII and HTML files. However, Adobe Acrobat PDF and Microsoft PowerPoint file formats are not supported at this time.

ZyIndex builds an index based on the documents specified and does not use the documents themselves for retrieval. However, as documents are added or changed, the index is automatically updated. Indexes created by ZyIndex can be very large--up to ten gigabytes can be indexed, or the equivalent of 100 gigabytes of documents. (Indexes of ten gigabytes normally represent approximately 100 gigabytes worth of documents.) If you need to restrict access to certain documents or indexes, ZyIndex allows you to define users and passwords that can be assigned to specific documents, groups of documents, or indexes in order to control security.

Because all indexing done by ZyIndex is on the complete text of the document, and indexes can be very large, a strong set of searching features is needed to ensure accuracy and relevancy of retrieval. ZyIndex Webserver supports Boolean operators and full nesting, phrase searching, advanced proximity searching and truncation, "fuzzy" searches that retrieve words similar to those specified, a "vocabulary" or browse index feature, field searching, and thesaurus for location of synonyms. Searching for numbers or number ranges is supported using standard math operators such as (is less than), (is greater than), =, etc.

In addition, the Concept feature allows web site managers to define searches that cover a particular subject contained in the index, name and save the search strategy, and then display the stored Concept searches for use by end-users searching the index. All of these features are included in an easy-to-use HTML template included with the package.

Retrieved documents, ranked according to relevancy, are automatically translated to HTML on-the-fly for viewing through web browsers regardless of native format, and can then be viewed in native format by launching the appropriate application program. Search terms are highlighted within the context of retrieved documents and users can move "hit to hit" to each

occurrence of the term(s) specified in their search request. Another nice feature of the ZyImage Webserver is the ability to view TIFF images directly through the web browser without the use of a helper application or plug-in, by using a TIFF to GIF converter included with the product.

From the ZyLab home page, you can view a demo of ZyIndex Webserver in action on a test database provided by the National Library of Medicine, as well as use a test database set up by ZyLabs to demonstrate a basic installation using the default searching interface and features.

CHOOSING THE RIGHT INTRANET SEARCH ENGINE FOR THE JOB

As demonstrated by the products reviewed in this article, there are many different things that need to be taken into consideration when evaluating and selecting a search engine for an intranet. Size of the site, the type of documents included, the number of web servers, server platform, and technical expertise available are all major factors influencing the selection of an intranet search engine.

If the intranet site is small and does not contain documents in formats other than HTML and ASCII text, the freeware search engines may be enough to do the job. The frequent downside to these free tools, however, is that advanced technical knowledge is needed to configure and customize the software for site-specific use, and that advanced searching functionality found in the commercial engines is not available. In addition, little formal technical support is offered by any of the free intranet search engines, except for Excite, which charges for its support and maintenance contract.

For large, highly-developed intranet sites, spending the money on a investment. Having the ability to index documents in a variety of file types, including distributed relational databases, and from a variety of locations, both internal and external, makes integrating and then retrieving information on an intranet much easier. Advanced searching features such as field, proximity, and concept searching, as well as the intelligent alerting capabilities promised with the next release of Verity's SEARCH'97 Agent Server, can reduce the number of irrelevant hits produced by a search of a large document collection and automate the search process so that users are automatically notified when content that matches their search profile is added.

WHAT WILL THE FUTURE HOLD?

In this fast-paced, ever-changing world of web-based information retrieval, there are several trends that promise to have a large effect on search and retrieval functionality of intranets.

Not Just Documents

Integration of access to distributed databases (not just documents) with intranet search engines is of paramount importance if intranets are to evolve to the next level of importance in the enterprise.

There are a host of vendors that provide gateways and development tools that can make access to distributed databases from the World Wide Web a reality. Intranet search engine vendors such as Fulcrum, Verity, and Open Text are poised to move to that next level, and may emerge as the favorites for intranet search engines in the near future.

Bundling With Server Software

More and more, web server software packages designed for intranet use are coming bundled with search engines designed to work with the web server software. The two major commercial web server vendors, Netscape and Microsoft, have already capitalized on this trend by including search engines as part of their web server offerings.

Netscape's Enterprise Server comes with the option of purchasing Verity's search engine and using Netscape's Catalog Server (based on Harvest) for indexing document collections. Microsoft's Internet

Information Server provides searching functionality through the Microsoft Index Server, a free package that can index and search HTML and file formats created by the software packages in the Microsoft Office Suite. Although both Netscape's Enterprise Server and Microsoft's Internet Information Server provide a built-in searching solution, other search engine products can be used with these web servers, if desired. As intranets grow, it is likely that even though a basic search engine may be included with a web server product, a separate search engine may be purchased as well, depending on the size and complexity of the intranet site.

Intelligent Agents on Alert

The addition of intelligent agents that can "remember" a search query and run it unattended against both internal and external indexes is another emerging trend that will surely become a favorite with intranet users. Products such as Verity's SEARCH'97 Agent Server and other products mentioned in the sidebar automate the search process and provide vital alerting services that can keep users up-to-date on topics in their areas of interest in "real-time" fashion. The personal search agent products also have the potential, when used with an intranet search engine, to combine results from the inside intranet world with the outside Internet world, giving users a comprehensive view of very current information on specified topics.

Getting It All

Will we ever really be able or even want to search all internal information and external information using one package? Already, typical end-users are becoming frustrated with the amount of retrieval returned by the popular Internet search engines. Intranets, as they grow, have the potential to inspire that same frustration if proper indexing and search and retrieval tools are not developed and implemented.

Gaining balance between providing relevancy, comprehensiveness, and manageability of information on intranets and the Internet as a whole, through development of a set of end-user tools for retrieving and filtering large sets of information, will provide one of the greatest challenges to information professionals in the coming months and years.

REFERENCES

(1) Levitt, Lee. "Intranets: Internet Technologies Deployed Behind the Firewall for Corporate Productivity." Prepared for the Internet Society INET '96 Annual Meeting. <<http://www.process.com/intranets/wp2.htm>>

(2) Balderston, Jim. "Search Engine Vendors Eye Intranets: Intranets Mean That Search Tools Must be Fine-Tuned for Corporate Needs." InfoWorld 18, No. 27 (July 1, 1996): pp. 41.

(3) Canada NewsWire. November 20, 1996.

Communications to the authors should be addressed to Peggy Zorn, Parke-Davis Pharmaceutical Research, 2800 Plymouth Road, Ann Arbor, MI 48150; 313/996-7202; zornm@aa.wl.com and/or Mary Emanoil, Parke-Davis Pharmaceutical Research, 2800 Plymouth Road, Ann Arbor, MI 48150; 313/996-1814; emanoim@aa.wl.com and/or Lucy Marshall, Edge Information Services, 2642 E. Cholla St., Phoenix, AZ 85028; 602/485-9363; edgeinfo@dancris.com and/or Mary Panek, United Technologies Research Center, 411 Silver Lane, MIS 129-01, East Hartford, CT 06108; 860/610-7478; panekmt@utcc.utc.com.

RELATED ARTICLE: The Next Level

SEARCHING DATABASES THROUGH AN INTRANET

While intranet search engines can index and search collections of documents on an intranet, what about including existing databases that a company might have, such as Oracle, Sybase, or Microsoft SQL Server databases? Often, the bulk of the most important information a company has

is stored in one of these formats. Gaining access to this vital information from a corporate intranet is a hot issue and will likely form the next wave of major intranet expansion.

Most intranet search engines are not yet able to integrate the searching of information to this next level: tapping both documents and collections of information stored in traditional database formats. Almost all major database vendors have created web-based interfaces and gateways to their products. There are many generic products available that allow you to interface to ODBC (open database connectivity)-compliant databases that use one development environment.

Below is a listing of major vendors competing in the expanding market of web-based connectivity to ODBC-compliant databases, with URLs for more information. Some, such as Oracle WebSystem and Sybase's NetImpact Dynamo, are targeted to specific database products. However, despite these specializations, all products still claim to have the ability to interface with any ODBC-compliant database product.

ColdFusion <http://www.allaire.com>

Everyware's Tango Enterprise <http://www.everyware.com/>

MEGASOFT Web Transporter <http://www.megasoft.com/>

Microsoft Internet Information Server with Microsoft dbWeb
<http://www.microsoft.com>

Netscape LiveWire <http://www.netscape.com>

NeXT WebObjects <http://www.next.com>

Oracle WebSystem <http://www.oracle.com.sg>

O'Reilly's WebSite Professional <http://www.ora.com/>

Sybase NetImpact Dynamo and web.sql <http://www.sybase.com>

WebDynamics Spider <http://www.w3spider.com>

RELATED ARTICLE: Intelligent Search Agents

Intelligent search agents allow users to create profiles based on their information needs and to simultaneously search selected sites from the external Web, corporate intranet, newsgroups, etc. for the desired information. It is similar to the use of alerting services or SDIs in traditional online searching, except that the intelligent agent can learn from the results, thereby refining the query and returning more valuable information with each new search.

The degree to which intelligent agents are being used varies among software products. Some are simply monitoring tools to alert users when changes have been made to bookmarked sites, but others make associations between search terms and other frequently occurring terms found in search results and then alert the user to these associations. Regardless of the level of agent sophistication, one can expect that software developers will continue to incorporate and improve upon this technology in their products.

Search software that currently uses agent technology includes CyberSearch and WebCompass. Frontier Technologies has announced the release of the 3.0 version of CyberSearch

(<http://www.frontiertech.com/products/cyberseb/csspecl.htm>), its Internet searching and bookmarking utility. Frontier calls the new version of CyberSearch "a global information management tool" because it searches documents on the Internet, intranet, and local PC. Through the use of standard Internet search engines such as Alta Vista, Lycos, Excite, and InfoSeek and server-side indexing of internal documents, this product incorporates the concept of seamless searching among all the information sources accessible to a user.

Quarterdeck Corporation intends to develop a version of its well-reviewed WebCompass software (<http://www.quarterdeck.com/qdeck/products/webcompass/>) that will not only allow users to query multiple search engines, as is the case with its

current release, but will also allow for the inclusion of intranet resources. The current version of WebCompass searches multiple Internet search engines simultaneously, sorts the results, and removes duplicate hits. Results are returned in a Microsoft Access database for easy manipulation.

Other software tools that incorporate intelligent agents and that may be beneficial for multisite searching are available. To keep abreast of new developments in the use of intelligent agents for intranet/Internet searching, visit the Complete Intranet Resource (<http://www.intrack.com/intranet/>). This site provides detailed information about intranets, including a list of software sources.

COPYRIGHT 1997 Online Inc.

SPECIAL FEATURES: table; illustration

DESCRIPTORS: Intranets--Usage; Online searching--Usage

PRODUCT/INDUSTRY NAMES: 7399200 (Info Services ex Database)

SIC CODES: 7389 Business services, not elsewhere classified

FILE SEGMENT: TI File 148